

Datenmodelle und Datenformate für die Modellierung des Fußballwortschatzes im Kicktionary

Thomas Schmidt thomas.schmidt@ids-mannheim.de Tel.: +49 621 1581-304

1. Einleitung

Das Kicktionary ist ein dreisprachiges (deutsch-englisch-französisches) elektronisches Wörterbuch der Fußballsprache. Es basiert auf einem Korpus von geschriebenen Fußballberichten und (in geringerem Umfang) gesprochenen Fußballkommentaren und nutzt die Ideen der Framesemantik (Fillmore 1982, Fillmore et al. 2003) sowie der lexikalischen Relationen (Fellbaum 1998) zur Strukturierung des Wortschatzes. Verschiedene Aspekte der Erstellung, Präsentation und Nutzung des Kicktionary sind in Schmidt (2008, 2009 und 2010) dargestellt. Im vorliegenden Beitrag konzentriere ich mich auf die Frage, welche Datenmodelle und welche Datenformate zur Modellierung des Wortschatzes im Kicktionary zum Einsatz kamen. Zu diesem Zweck möchte ich einleitend zunächst mein Verständnis dieser drei Begriffe – Datenmodell, Datenformat und Modellierung – näher erläutern.

Unter dem *Datenmodell* des Kicktionary verstehe ich die Struktur, mittels derer die Gesamtheit der Wörterbucheinheiten und deren Beziehungen untereinander erfasst werden. Das Datenmodell ist abstrakt, d.h., es lässt sich zwar (symbolisch/verbal) beschreiben, hat aber selbst keine konkrete physikalische Manifestation, sondern wird nur beispielsweise in Form des zugehörigen Datenformats oder der nutzerorientierten Visualisierung des Wörterbuchs direkt erfassbar. Das *Datenformat* ist in diesem Sinne also die (oder eine) konkrete physikalische Repräsentation des Datenmodells. Es hat typischerweise die Form einer oder mehrerer XML-Dateien oder einer relationalen Datenbank. Mit dieser Unterscheidung folge ich der in der Datenbanktechnologie üblichen Differenzierung zwischen physikalischer und logischer Ebene der Datenverarbeitung, der als drittes die Anwendungsebene zur Seite gestellt wird (vgl. z.B. Date 1995). Unter *Modellierung* schließlich verstehe ich den empirischen Prozess, in dem das Modell aufgebaut, ggf. modifiziert und mit Inhalt gefüllt wird. Konkreter: Unter den Begriff der Modellierung fasse ich all diejenigen Vorgänge in der Wörterbucherstellung, in denen Wörterbucheinträge erstellt, ergänzt und geändert und in die (sich eventuell ebenfalls ändernde) Struktur des Wörterbuchs eingeordnet werden. Das Datenmodell, das in einem Datenformat physikalisch repräsentiert ist, ist in diesem Sinne das *Produkt*, die Modellierung der *Prozess* der Wörterbucherstellung.³³

Im folgenden Abschnitt beschreibe ich also zunächst Datenmodell und -format(e) des Kicktionary. Anschließend stelle ich den Modellierungsprozess und einige damit zusammenhängende Schwierigkeiten dar. Schließlich skizziere ich zwei Änderungen am Modell bzw. am Modellierungsprozess, die mir geeignet scheinen, zumindest einen Teil dieser Schwierigkeiten zu überwinden.

³³ Ich wurde zu Recht darauf hingewiesen, dass dieses Verständnis des Begriffs „Modellierung“ weiter gefasst ist als in der Metalexikographie üblich – vgl. z.B. Geeb (2001), S. 29: „Datenmodellierung sei verstanden als der Entwurf von (lexikographischen) Dateneinheiten und ihren funktionalen Beziehungen ohne Rücksichtnahme auf bereits bestehendes Datenmaterial.“ Wie ich weiter unten darlege, ist es aber gerade die Wechselwirkung vorgegebener Datenstrukturen mit ebenjenem „bereits bestehenden [oder neu hinzukommendem] Datenmaterial“, die sich bei der Erstellung des Kicktionary als interessantes Problem erwiesen hat. Vor diesem Hintergrund erscheint mir die Erweiterung der Bedeutung des Begriffs „Modellierung“ für die Zwecke dieses Beitrags zu rechtfertigen.

2. Datenmodell und Datenformate

Zentrale Einheit im Datenmodell des Kicktionary ist – wie im Berkeley FrameNet – die LEXIKALISCHE EINHEIT (abgekürzt: LU), die Ruppenhofer et al. (2010, S. 5) als „a pairing of a word form with a meaning“ definieren.³⁴ Die lexikalische Einheit ist insofern zentral, als sich jeder andere Modellbestandteil zu ihr in eine direkte Beziehung setzen lässt – jede andere Wörterbucheinheit ist entweder Bestandteil einer lexikalischen Einheit oder dient dazu, lexikalische Einheiten zueinander in Beziehung zu setzen oder in größere Struktureinheiten zusammenzufassen.

Zu einer lexikalischen Einheit gehören erstens (obligatorisch) Angaben zur Sprache und zum PART-OF-SPEECH, zweitens (optional) eine DEFINITION und drittens (obligatorisch) eine Reihe von KORPUSBEISPIELEN, in denen (u.a.) die verwendete Form der lexikalischen Einheit markiert ist.

Lexikalische Einheit:	<i>tunneln</i>
Sprache:	Deutsch
Part-Of-Speech:	Verb
Definition:	Der ballführende Spieler behauptet den Ballbesitz gegen einen angreifenden Gegenspieler, indem er ihm den Ball zwischen den Beinen hindurch spielt.
Korpusbeispiele:	
(1)	Diogo Rincón tunnelte Paul Freier im Strafraum und sein Schuss trudelte [...] an Jörg Butt vorbei und landete in Netz.
(2)	[...] nur wenige Sekunden später war es Duff, der nach einer herrlichen Kombination zwischen Kezman und Cole [...] den spanischen Keeper zum 3:0 tunnelte .
(3)	Ailton tunnelte Chris an der Strafraumgrenze und spielte so Klasnic frei.

Abb. 1: Beispiel für eine lexikalische Einheit mit Korpusbeispielen, vgl. http://www.kicktionary.de/LUs/Beat/LU_391.html

Jede lexikalische Einheit ist genau einem FRAME zugeordnet, der seinerseits Bestandteil genau einer SCENE ist. Eine Scene wird dabei verstanden als das (u.U. nonverbale) Wissen über einen prototypischen Handlungsablauf; ein Frame fasst alle lexikalischen Einheiten zusammen, mittels derer ein bestimmter Aspekt dieses Handlungsablaufes aus einer bestimmten Perspektive sprachlich dargestellt werden kann. Zu einer Scene gehört eine Charakterisierung des Handlungsablaufes selbst sowie der typischerweise an ihr beteiligten Aktanten und Gegenstände. Letztere werden FRAME-ELEMENTE genannt.

³⁴ Eine Form korrespondiert hier also immer nur mit einer Einzelbedeutung. Eine Einheit wie das „Stichwort“ in einem klassischen Wörterbuch, unter der mehrere Einzelbedeutungen zusammengefasst werden, gibt es beim Kicktionary nicht als explizite Einheit im Modell.

Scene:	One-on-One
Handlungsablauf:	Der ballführende Spieler wird von einem Gegenspieler, der versucht, in Ballbesitz zu gelangen, angegriffen. Ein Zweikampf endet, indem entweder der ballführende Spieler den Ballbesitz behauptet oder der angreifende Spieler in Ballbesitz gelangt.
Frame-Elemente:	BALLFÜHRENDER_SPIELER, ANGREIFENDER_SPIELER, BALL
Zugehörige Frames:	Challenge (LUs: <i>angreifen, attackieren, bedrängen, stören</i> etc.) Beat (LUs: <i>ausspielen, umdribbeln, vernaschen, tunneln</i> etc.) Deny (LUs: <i>abgrätschen, abjagen, blocken, stoppen</i> etc.)

Abb. 2: Beispiel für eine Scene mit zugehörigen Frames, vgl. http://www.kicktionary.de/One_On_One_Scenario.html

Die Frame-Elemente finden sich in den Korpusbeispielen als Annotationen der Mitspieler der lexikalischen Einheit wieder.

(1)	[Diogo Rincón] _{BALLFÜHRENDER_SPIELER} tunnelte [Paul Freier] _{ANGREIFENDER_SPIELER} im Strafraum und sein Schuss trudelte [...] an Jörg Butt vorbei und landete in Netz.
(2)	[...] nur wenige Sekunden später war es [Duff] _{BALLFÜHRENDER_SPIELER} , der nach einer herrlichen Kombination zwischen Kezman und Cole [...] [den spanischen Keeper] _{ANGREIFENDER_SPIELER} zum 3:0 tunnelte .
(3)	[Ailton] _{BALLFÜHRENDER_SPIELER} tunnelte [Chris] _{ANGREIFENDER_SPIELER} an der Strafraumgrenze und spielte so Klasnic frei.

Abb. 3: Korpusbeispiele mit annotierten Frame-Elementen

Die einzelnen Wörterbucheinträge – lexikalische Einheiten mit den zugehörigen Spezifizierungen und Beispielen – werden also einerseits in einer Hierarchie aus Scenes und Frames organisiert.

Andererseits werden, dem Ansatz von WordNet folgend, bedeutungsgleiche lexikalische Einheiten zu SYNSETS zusammengefasst, wobei das Kicktionary als mehrsprachiges Wörterbuch auch die Übersetzungsäquivalenz als eine Form der Synonymie behandelt. Zwischen Synsets können darüber hinaus die SEMANTISCHEN RELATIONEN Hyperonymie/Hyperonymie (Ober- bzw. Unterbegriff bei Substantiven), Meronymie/Holonymie (Teil-Ganzes-Beziehung bei Substantiven) sowie Troponymie (Analogon zur Hyperonymie bei Verben) im Datenmodell festgehalten werden.

Synset:	{ <i>tunneln / to nutmeg</i> }
Troponyme:	{ <i>ausspielen, sich durchsetzen / beat, round / éliminer</i> }
Synset:	{ <i>Innenverteidiger / central defender, centre-back / défenseur central</i> }
Hyperonyme:	{ <i>Verteidiger, Abwehrspieler / defender / arrière, défenseur</i> }
Holonyme:	{ <i>Verteidigung, Abwehr / defence, backline / arrièregarde, défense</i> }

Abb. 4: Beispiele für Synsets und semantische Relationen

Durch Verkettung der transitiven semantischen Relationen entstehen weitere hierarchische Organisationen des Wortschatzes, die im Kicktionary KONZEPHTHIERARCHIEN genannt werden.

Mannschaft, Team / side, squad / équipe, formation
 Schlussmann, Torhüter / custodian, goalkeeper / gardien, portier
 Abwehr, Verteidigung / backline, defence / arrière-garde, défense
 Innenverteidigung / central defence / défense centrale
 Innenverteidiger / central defender / défenseur central
 Mittelfeld / midfield / milieu de terrain
 Mittelfeldspieler / midfielder / milieu
 Angriff, Sturm / attack / attaque
 Mittelstürmer / centre-forward / avant-centre

Abb. 5: Beispiel für eine Konzepthierarchie, basierend auf der semantischen Relation Meronymie/Holonymie, vgl. http://www.kicktionary.de/CONCEPT_HIERARCHIES/Hyponymy_Individual_Actors.html

Anders als bei der Scenes-Frames-Hierarchie gilt nicht, dass eine lexikalische Einheit genau einer Konzepthierarchie zugeordnet sein muss. Es tritt sowohl der Fall auf, dass eine lexikalische Einheit gar nicht in einer solchen Hierarchie auftaucht (weil es keine weiteren lexikalischen Einheiten gibt, mit denen sie in einer semantischen Relation steht), als auch der Fall, dass sie Bestandteil mehrerer solcher Hierarchien ist (weil sie z.B. sowohl Holonym als auch Hyperonym von anderen lexikalischen Einheiten ist).

Abbildung 6 stellt noch einmal die wichtigsten Einheiten im Kicktionary-Datenmodell und ihre Beziehungen zueinander dar.

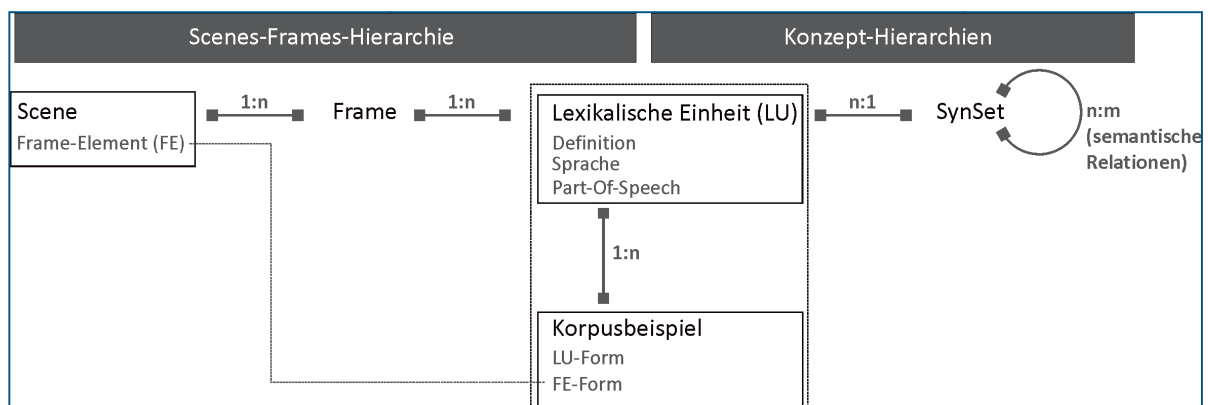


Abb. 6: Schema des Kicktionary-Datenmodells

Auf der physikalischen Ebene wurde dieses Modell während der Erstellung in Form einer einfachen XML-Datei repräsentiert, wie in Abbildung 7 illustriert. Da das Kicktionary von vornherein als ein zeitlich befristetes Projekt mit einem einzigen Bearbeiter angelegt war, werden nicht alle strukturellen Beschränkungen des Modells auf der Ebene des Dateiformats abgebildet, d.h., die zugehörige DTD stellt zwar z.B. sicher, dass ID-Bezüge in den Frame-Element-Annotationen (Attribute `@fe-idref`) auf eine existierende ID (Attribute `@name` in den Elementen `<argument>`) verweisen, nicht aber, dass dieser Bezug sich auch innerhalb der betreffenden lexikalischen Einheit befindet oder über verschiedene LUs innerhalb eines Frames konsistent benannt ist. Die Integrität des Datenformats im Bezug auf das Modell wurde stattdessen teilweise durch eine entsprechende Funktionalität in der Bearbeitungssoftware (ein eigens für das Projekt entwickelter Editor mit Konkordanzfunktion für das Korpus, siehe Schmidt 2008), teilweise einfach durch regelmäßige manuelle Kontrollen sichergestellt.

```

<LU_REPORT id="LU_391">
<LEXICAL-UNIT scenario="One_On_One" frame="Beat"
    lu-id="tunneln.v" lang="de" wordclass="v" synset="TO_NUTMEG">tunneln</LEXICAL-
UNIT>
<arguments>
<argument name="PLAYER_WITH_BALL"/>
<argument name="AREA"/>
<argument name="OPPONENT_PLAYER"/>
</arguments>
<DEFINITION>Der ballführende Spieler (PLAYER) überwindet einen angreifenden Gegenspieler
(OPPONENT_PLAYER), indem er ihm den Ball zwischen den Beinen hindurch spielt. Der
Ort, an dem dieser Zweikampf stattfindet (LOCATION) kann erwähnt werden.
</DEFINITION>
<EXAMPLE lang="de">
<FE_REF fe-idref="PLAYER_WITH_BALL">Diogo Rincón</FE_REF>
<LU_REF lu-idref="tunneln.v">tunnelte</LU_REF>
<FE_REF fe-idref="OPPONENT_PLAYER">Paul Freier</FE_REF>
<FE_REF fe-idref="AREA">im Strafraum</FE_REF>
    und sein Schuss trudelte, begünstigt durch
    einen Platzfehler, an Jörg Butt vorbei und landete in Netz.
<EXAMPLE lang="de" source-element-id="p5" source-text-id="K_175c">
<FE_REF fe-idref="PLAYER_WITH_BALL">Ailton</FE_REF>
<LU_REF lu-idref="tunneln.v">tunnelte</LU_REF>
<FE_REF fe-idref="OPPONENT_PLAYER">Chris</FE_REF>
<FE_REF fe-idref="AREA">an der Strafraumgrenze</FE_REF> und spielte so Klasnic frei.
</EXAMPLE>
</LU_REPORT>

```

Abb. 7: XML-Repräsentation einer lexikalischen Einheit

3. Modellierung

Ausgangspunkt für die Modellierung des Fußballwortschatzes im Kicktionary waren einfache Wortlisten, die aus den Korpora der drei beteiligten Sprachen automatisch generiert und nach Häufigkeit der Vorkommen geordnet wurden. Auf der Grundlage dieser Wortlisten wurde dann für jede Sprache eine Liste von je ca. 100 lexikalischen Einheiten erstellt, die entweder spezifisch für die Fußballsprache sind oder zumindest in Fußballberichten in deutlich erhöhter Frequenz auftreten. Dieses Material bildete dann die Basis für die weitergehenden Analysen.

Das weitere Vorgehen zum Aufbau der Scenes-Frames-Hierarchie bestand in einem iterativen Prozess: Das vorhandene Material lexikalischer Einheiten wurde zunächst anhand von Vorkommen im Korpus analysiert und nach Ähnlichkeiten im Bezug auf Argumentstrukturen und semantische Eigenschaften geordnet. Daraufhin wurden erste Scenes und Frames mit den zugehörigen Frame-Elementen definiert, für jede lexikalische Einheit eine Zuordnung zu einem Frame vorgenommen und Korpusbeispiele gemäß den betreffenden Frame-Elementen annotiert.

Meine Grundannahme war dabei, dass es nicht darum geht, *die* „korrekten“ Scenes und Frames der Fußballsprache zu ermitteln und zu beschreiben. Vielmehr gehe ich davon aus, dass es sich bei der Scenes- und Frames-Analyse um ein wissenschaftliches Modell handelt, das nicht an sich falsch oder richtig sein kann. Sein Wert bemisst sich eher danach, wie nützlich und handhabbar seine verkürzende Abbildung der Realität im Bezug auf einen bestimmten Erkenntnis- oder Anwendungszweck ist (vgl. dazu die Ausführungen zum pragmatischen Merkmal wissenschaftlicher Modelle in Stachowiak 1973). In diesem Sinne begreife ich Scenes und Frames in erster Linie als Mittel zur Makrostrukturierung eines Wörterbuchs (und nicht z.B. als Ansatz zur Beschreibung oder Erklärung kognitiver Strukturen oder als Basis für eine maschinelle Sprachverarbeitung). Ihr Sinn besteht für mich darin, die potentiell unüberschaubare Gesamtheit lexikalischer Einheiten dergestalt in kleinere Einheiten zu strukturieren, dass es dem Benutzer ermöglicht bzw. vereinfacht wird, Beziehungen im Wortschatz

zu erkennen und für die Sprachrezeption oder -produktion fruchtbar zu machen. Nützliche Frames sind demnach solche, die

- erstens eine hilfreiche „Portionierung“ des Gesamtwortschatzes leisten. Frames, die zu viele lexikalische Einheiten enthalten, sind nach diesem Kriterium ebenso defizitär wie Frames, die zu wenige lexikalische Einheiten enthalten. Es sei zur Klarstellung angemerkt, dass die Anzahl der Mitglieder eines Frames vor allem mit seinem Abstraktionsgrad oder dem der übergeordneten Scene variiert, die Framesemantik aber keine Aussagen über einen angemessenen Abstraktionsgrad macht.
- zweitens nur solche lexikalischen Einheiten enthalten, die bezüglich ihrer Semantik und Argumentstrukturen ausreichend homogen sind. Eine zu geringe Homogenität äußert sich beispielsweise darin, dass die annotierten Korpusbeispiele einzelner LUs eine sehr stark variierende Auswahl aus den zur Verfügung stehenden Frame-Elementen treffen.

Während der Modellierung wurde also in regelmäßigen Abständen geprüft, ob die entstehende Scenes-Frames-Hierarchie hinsichtlich dieser beiden Kriterien und aus Sicht eines Wörterbuchbenutzers als ausreichend „nützlich“ und „handhabbar“ zu beurteilen ist.³⁵ War dies nicht der Fall, wurde die vorhandene Struktur modifiziert (gelegentlich auch vollständig verworfen), indem beispielsweise zu umfangreiche Scenes und Frames in mehrere unterteilt oder zu wenig umfangreiche in größere zusammengefasst wurden. Die abstrakten strukturellen Vorgaben des Datenmodells (wie in Abbildung 6 dargestellt) wurden in diesem Prozess allerdings nicht verändert – es ging lediglich darum auszuloten, wie die Freiräume in diesen Vorgaben (etwa: der Abstraktionsgrad einer Scene) optimal zu nutzen sind.

4. Schwierigkeiten bei der Modellierung

Obwohl die im vorigen Abschnitt beschriebene Form der Modellierung sich insgesamt als praktikabel erwiesen hat, haben sich während der Arbeit am Kicktionary dennoch einige Schwächen offenbart. Teilweise liegen diese darin begründet, dass die Methode der Frameanalyse – auch in den Arbeiten Fillmores – in aller Regel lediglich exemplarisch illustriert, nicht aber in Form allgemeingültiger Prinzipien für ein empirisches Arbeiten ausformuliert ist. Die konkrete Vorgehensweise beim Erstellen von Frames und die Beurteilung ihrer „Korrektheit“ bzw. Adäquatheit bleiben daher zu einem großen Teil im Ermessen des Lexikographen, der das empirische Material bearbeitet, und lassen sich kaum durch aus der Theorie abgeleitete Kriterien absichern.

Für die Schwierigkeiten in der Modellierung des Kicktionary war vor allem der Umstand verantwortlich, dass neu hinzukommendes Material (neue lexikalische Einheiten oder neue Beispiele aus dem Korpus) immer wieder die bestehende Scenes- und Frames-Struktur in Frage gestellt hat. Lokale Ergänzungen oder Modifikationen machten somit häufig Änderungen an der globalen Struktur des Wörterbuchs notwendig, deren Auswirkungen sich nur schwer abschätzen ließen, und die in jedem Falle mühsam und zeitaufwendig umzusetzen waren. In einer Diskussion auf der Lexicography-Mailingliste³⁶ hat Patrick Hanks die Ursache dieses Problems treffend wie folgt beschrieben: „FrameNet proceeds frame by frame, not word by

³⁵ Mir ist bewusst, dass dies keine im wissenschaftlichen Sinne ausreichend objektivierbaren Begriffe sind. Ob und wie eine solche Objektivierung zu leisten ist (beispielsweise im Rahmen einer wissenschaftlich fundierten Wörterbuchkritik), möchte und kann ich im Rahmen dieses Beitrags nicht diskutieren.

³⁶ <http://tech.groups.yahoo.com/group/lexicographylist/message/3178>

word. This may seem a trivial point, but it isn't. Although FrameNet uses empirical data, it does not use an empirical methodology.“

Der oben beschriebene Modellierungsprozess des Kicktionary folgt der hier kritisierten Methode von FrameNet insofern, als er Frames bzw. Scenes vordefiniert und das empirische Material dann in diese Struktur einzufügen versucht. Im Gegensatz zu FrameNet wurde beim Kicktionary aber dennoch versucht, den gesamten relevanten Wortschatz, so wie er sich im Korpus auffinden und belegen lässt, abzudecken. Statt Wörter oder Korpusbeispiele, die nicht in die vorhandene Struktur passen, zunächst oder dauerhaft auszusortieren, musste daher die Struktur selbst fortwährend modifiziert werden, was zu den oben geschilderten Problemen führte.

Um diesen Problemen zu begegnen, scheinen mir zwei Änderungen am Modellierungsprozess bzw. am Modell selbst erforderlich zu sein:

Erstens sollten die Beschreibung einzelner lexikalischer Einheiten und die Annotation zugehöriger Korpusbeispiele nicht von einer übergeordneten Scenes- und Frames-Struktur abhängig sein. Stattdessen sollte sich die Annotation der Argumentstruktur einer lexikalischen Einheit zunächst ausschließlich an der empirischen Analyse der lexikalischen Einheit selbst orientieren. Es ist zu erwarten, dass die auf diese Weise entstehenden Annotationen weniger anfällig für Neu- oder Uminterpretationen sind, einfach weil sie weniger interpretationsabhängige Abstraktion beinhalten. Statt also die Scenes- und Frames-Struktur in einer Top-down-Richtung dem empirisch untersuchten Material aufzuzwängen (bzw. sie in einem schwer zu kontrollierenden Wechselspiel mit der Empirie fortwährend zu modifizieren), könnte und sollte sie in einer Bottom-up-Richtung aus einer Bearbeitung des empirischen Materials auf einer weniger abstrakten Ebene abgeleitet werden. Das Modell selbst würde sich dadurch nicht grundlegend ändern, wohl aber der Modellierungsprozess – die Annotation von Korpusbeispielen würde der Konstruktion der Scenes- und Frames-Hierarchie nämlich dann eindeutig vorausgehen.

Zweitens hat die Arbeit am Kicktionary auch deutlich gemacht, dass Scenes- und Frames-Analysen nicht unbedingt für den gesamten Wortschatz eine gleichermaßen hilfreiche Methode darstellen. Scenes und Frames sind vor allem deshalb nützliche Elemente zur Strukturierung des Wortschatzes, weil sich über sie das (teilweise nichtsprachliche) Wissen über prototypische Handlungsabläufe in Bezug zu sprachlichen Mitteln setzen lässt, mit denen dieses Wissen ausgedrückt werden kann. Große Bereiche der Fußballsprache, wie zum Beispiel die Terminologie für verschiedene Spielerpositionen (*Libero*, *Mittelfeldregisseur*, *Außenstürmer* etc.) oder für Dinge auf dem Spielfeld (*Torpfosten*, *Eckfahne*, *Mittellinie*, *Strafraum* etc.) beschreiben aber gerade keine dynamischen Handlungen, sondern mehr oder weniger statische Objekte. In diesen Bereichen des Wortschatzes sind daher semantische Relationen ein deutlich nützlicheres und auch intuitiver anwendbares Mittel der Wörterbuchstrukturierung als Scenes und Frames. Fillmore (1978, S. 264) selbst erkennt dies auch an:

I think that semantic theory must reject the suggestion that all meanings need to be described in the same terms. I think, in fact, that semantic domains are going to differ from each other according to the kind of 'definitional base' which is most appropriate to them.

Die Entscheidung, im Modell des Kicktionary eine Zuordnung jeder lexikalischen Einheit zu (genau) einem Frame zu fordern, hat sich vor diesem Hintergrund als nicht sinnvoll erwiesen. Wenn Scenes und Frames ohnehin – wie oben beschrieben – erst entwickelt und definiert werden, nachdem die Beschreibung einzelner lexikalischer Einheiten mit annotierten Korpus-

beispielen abgeschlossen ist, besteht auch keine Notwendigkeit mehr, eine solche Zuordnung zu erzwingen. Die Scenes- und Frames-Hierarchie könnte und sollte sich dann auf denjenigen Bereich des Wortschatzes beschränken, in dem sie sich im Sinne des obigen Zitats als eine „appropriate definitional base“ erweist.

5. Literatur

- Date, Chris J. (1995): *An Introduction to Database Systems*. New York.
- Fellbaum, Christiane (Hg.) (1998): *WordNet – An Electronic Lexical Database*. Boston.
- Fillmore, Charles (1978): On the Organization of Semantic Information in the Lexicon. In: Farkas, Donka et al. (Hg.): *Papers from the Parasession on the Lexicon*, Chicago Linguistic Society, April 14-15, 1978. Reprint in: Fillmore, Charles: *Form and Meaning in Language: Volume I, Papers on Semantic Roles*. Stanford, S. 261-289.
- Fillmore, Charles (1982): Frame semantics. In: *Linguistics in the Morning Calm. Selected Papers from SICOL-1981*. Edited by the Linguistic Society of Korea. Seoul, S. 111-137.
- Fillmore, Charles/Johnson, Christopher R./Petruck, Miriam R. L. (2003): Background to Framenet. In: *International Journal of Lexicography* 16.3, S. 235-250.
- Geeb, Franziskus (2001): leXeML – Vorschlag und Diskussion einer (meta)-lexikographischen Auszeichnungssprache. In: *Sprache und Datenverarbeitung* 2, S. 27-61.
- Kicktionary – The multilingual electronic dictionary of football language: http://www.kicktionary.de/index_de.html (Stand: Oktober 2011).
- Ruppenhofer, Josef/Ellsworth, Michael/Petruck, Miriam/Johnson, Chris/Scheffczyk, Jan (2010): *FrameNet II: Extended Theory and Practice*. Internet: <http://framenet.icsi.berkeley.edu/book/book.html> (Stand: Oktober 2011).
- Schmidt, Thomas (2008): The Kicktionary Revisited. In: Storrer, Angelika/Geyken, Alexander/Siebert, Alexander/Würzner, Kay-Michael (Hg.): *Text Resources and Lexical Knowledge*. Berlin, S. 239-252.
- Schmidt, Thomas (2009): The Kicktionary – A Multilingual Lexical Resource of Football Language. In: Boas, Hans C. (Hg.): *Multilingual Framenets in Computational Lexicography*. New York, S. 101-134.
- Schmidt, Thomas (2010): Der Fußballwortschatz im Kicktionary. In: *Der Deutschunterricht* 3, S. 17-25.
- Stachowiak, Herbert (1973): *Allgemeine Modelltheorie*. Wien.